SPEECH
SYNTHESIS SYSTEM 100

TARGET WORD SEQUENCE /110
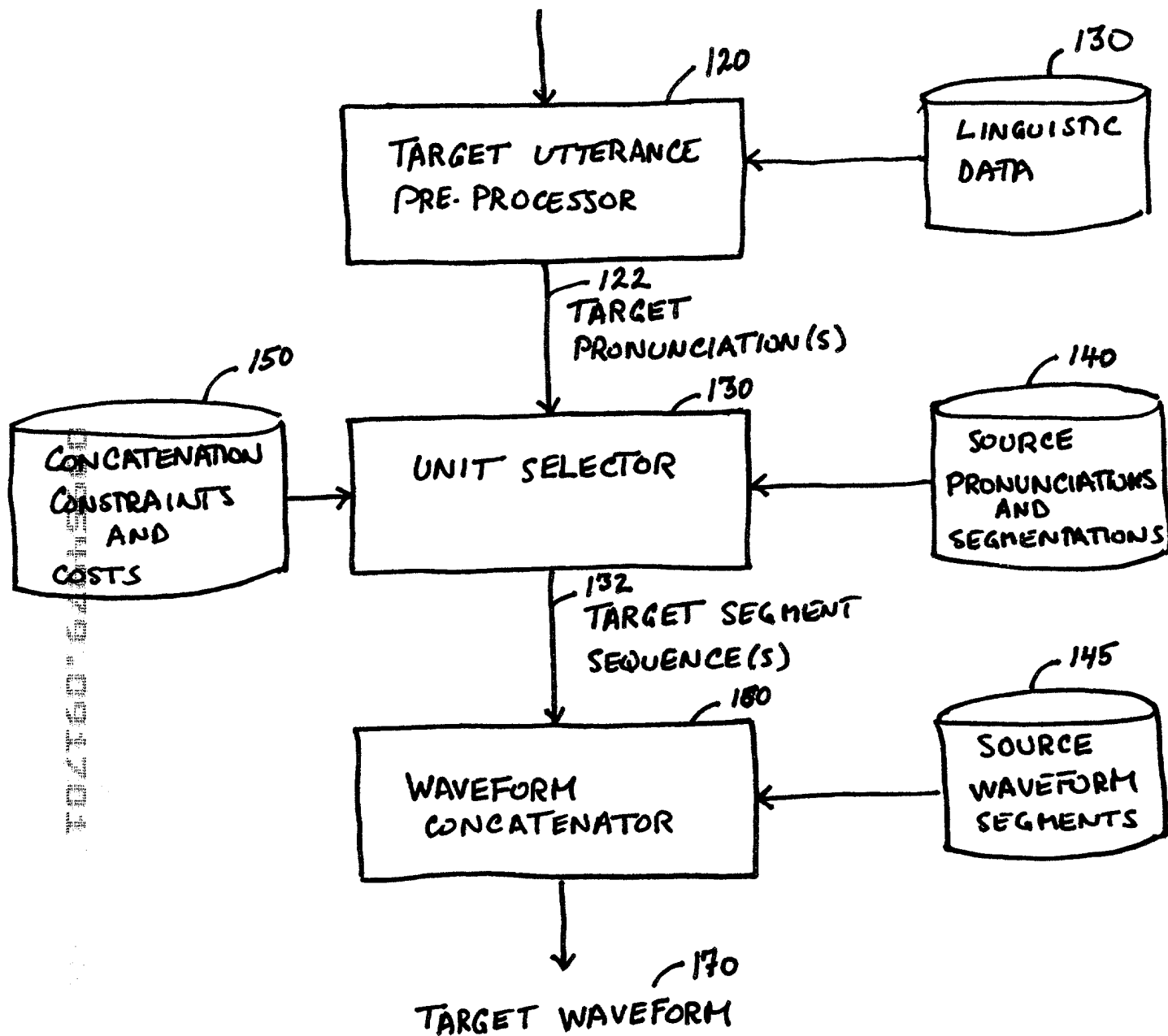
↓

TARGET UTTERANCE PRE-PROCESSOR /120 ← LINGUISTIC DATA /130
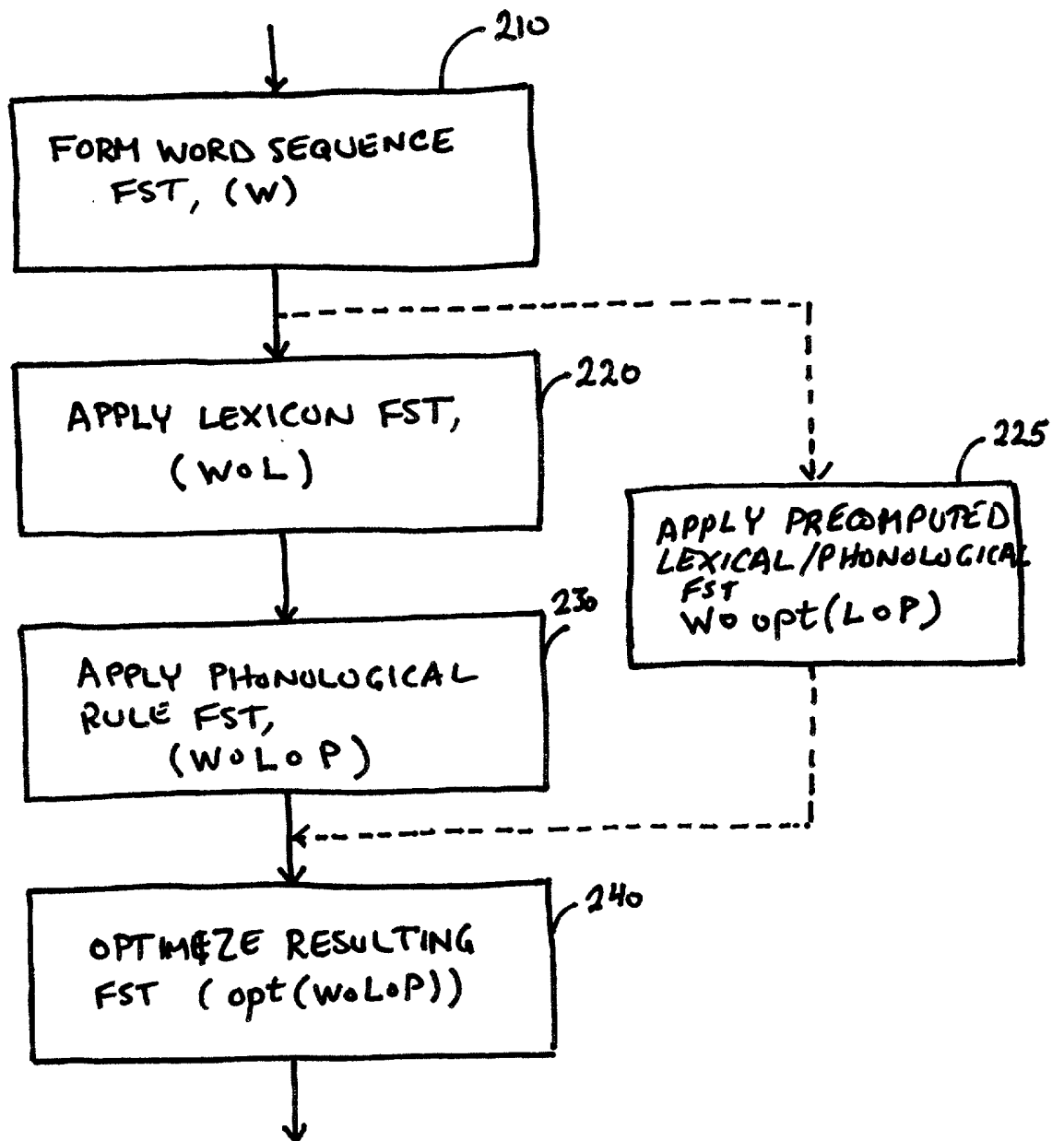
↓ 122
TARGET PRONUNCIATION(S)

CONCATENATION CONSTRAINTS AND COSTS /150 → UNIT SELECTOR /130 ← SOURCE PRONUNCIATIONS AND SEGMENTATIONS /140

↓ 132
TARGET SEGMENT SEQUENCE(S)

WAVEFORM CONCATENATOR /160 ← SOURCE WAVEFORM SEGMENTS /145

↓

TARGET WAVEFORM /170

FIG. 1

TARGET UTTERANCE PRE-PROCESSING:

FORM WORD SEQUENCE
FST, (W)   ~210

APPLY LEXICON FST,
(WoL)   ~220

APPLY PHONOLOGICAL
RULE FST,
(WoLoP)   ~230

APPLY PRECOMPUTED
LEXICAL/PHONOLOGICAL
FST
Wo opt(LoP)   ~225

OPTIMIZE RESULTING
FST (opt(WoLoP))   ~240

FIG. 2

UNIT SELECTION:

APPLY TRANSITION
FST, $((W \circ L \circ P) \circ T)$ — 310

APPLY SYNTHESIS FST
$(W \circ L \circ P \circ T) \circ S$ — 320

FIG. 3

FORMING SYNTHESIS FST (S)

FORM CONSTRAINT
  KERNEL   410

FORM PRONUNCIATION
GRAPH FOR EACH
  SOURCE UTTERANCE   420

LINK SOURCE
UTTERANCE GRAPHS
AND CONSTRAINT KERNEL   430

FIG. 4
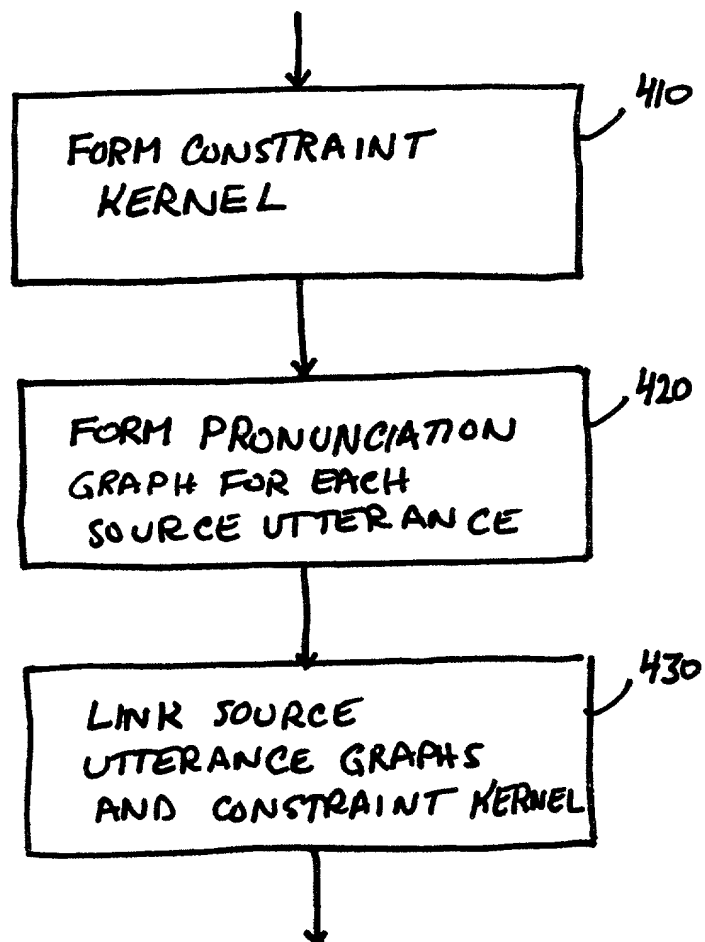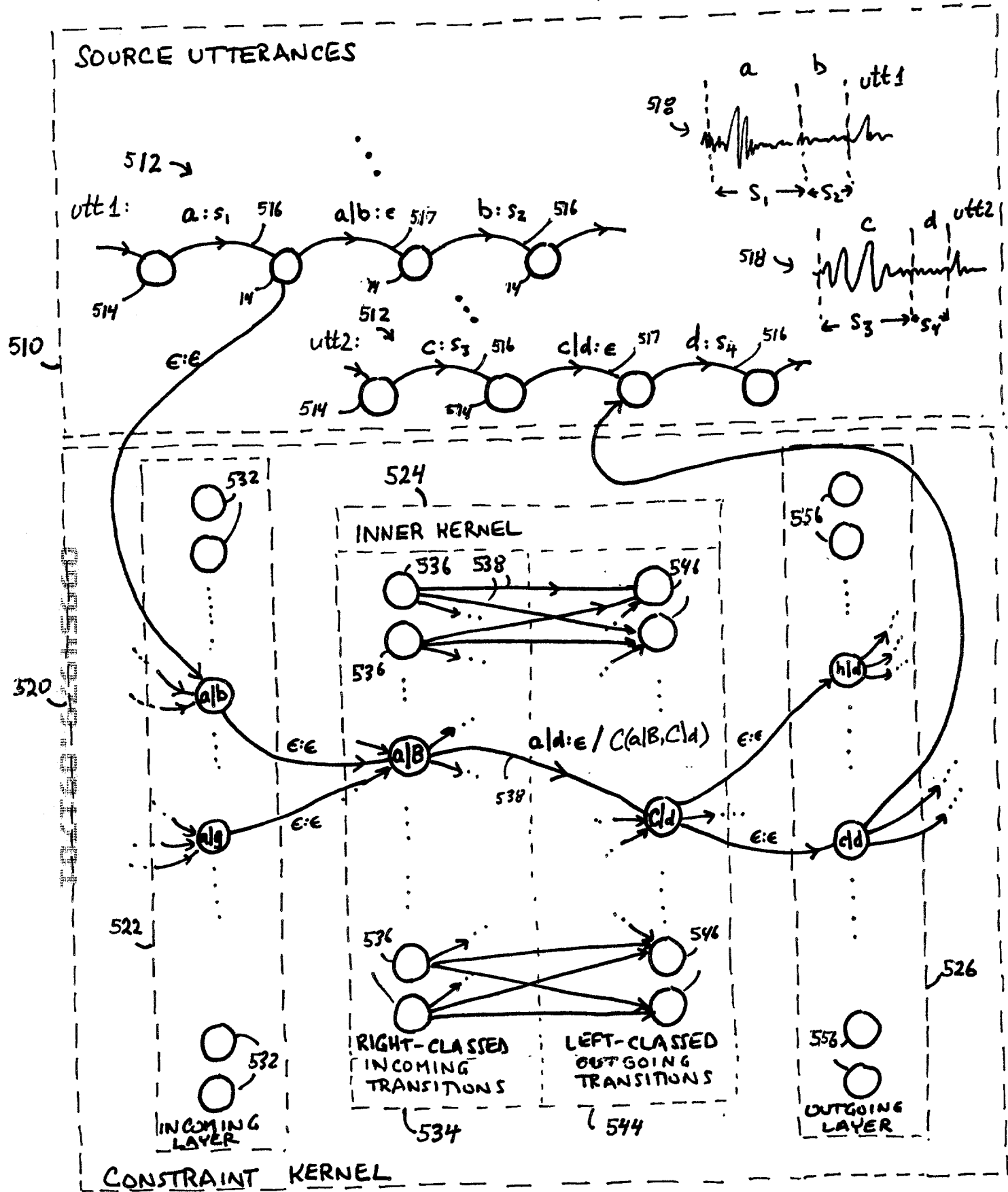
FIG.5

FIG. 6
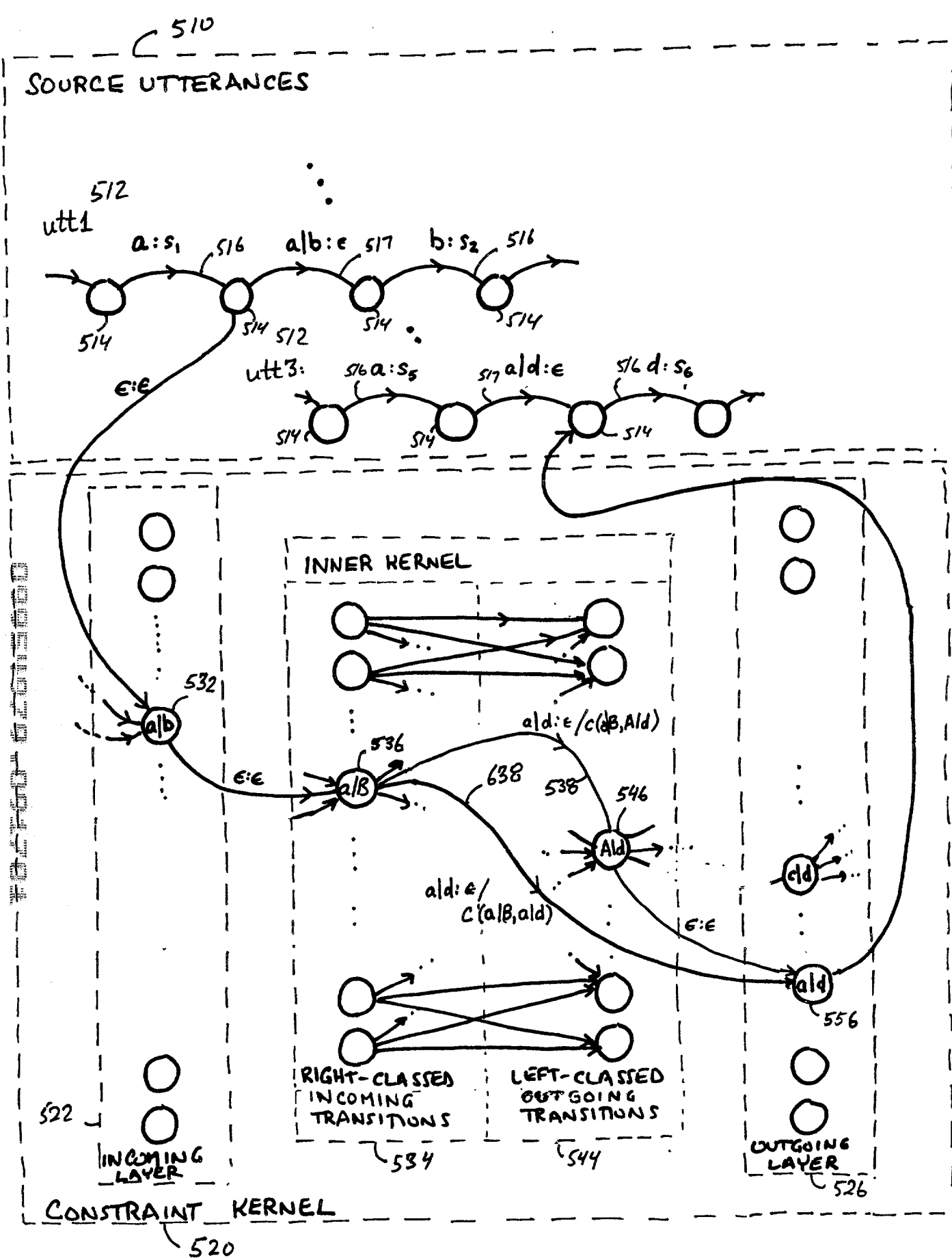
FIG. 7

SOURCE UTTERANCES

utt 1:   a:s₁   a|b:ε   b:s₂

ε:ε

utt5   a:s₅   a|b:ε   b:s₆

INNER KERNEL

S32

a|b   a|b:ε / c(a|b, a|b)   a|b
                                                                                                  556
                          ε:ε

a|a   ε:ε

838

556

RIGHT-CLASSED   LEFT-CLASSED
INCOMING   OUTGOING
TRANSITIONS   TRANSITIONS

522
INCOMING   OUTGOING
LAYER   LAYER   526
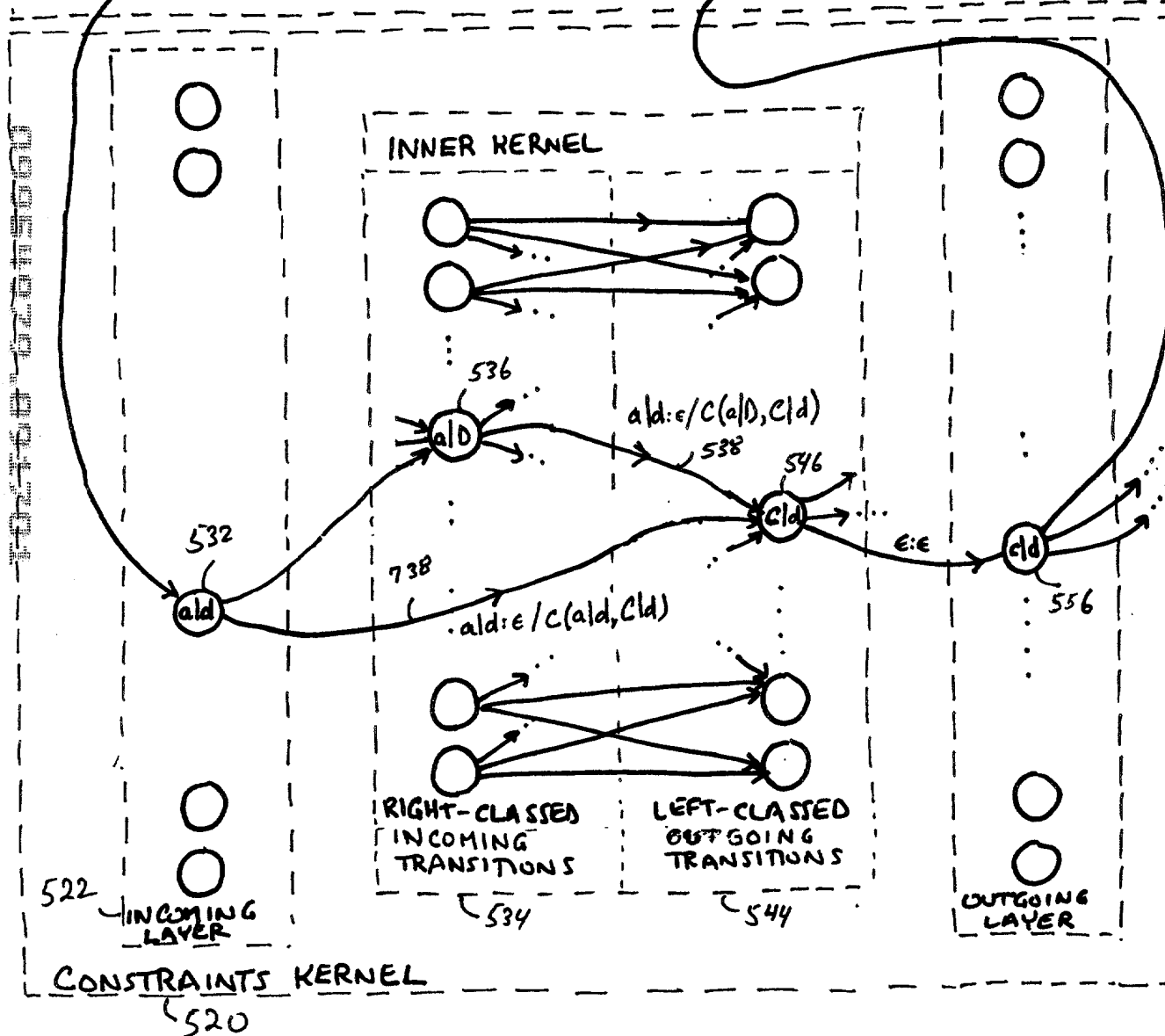
CONSTRAINTS KERNEL

520

FIG. 8

FIG. 9

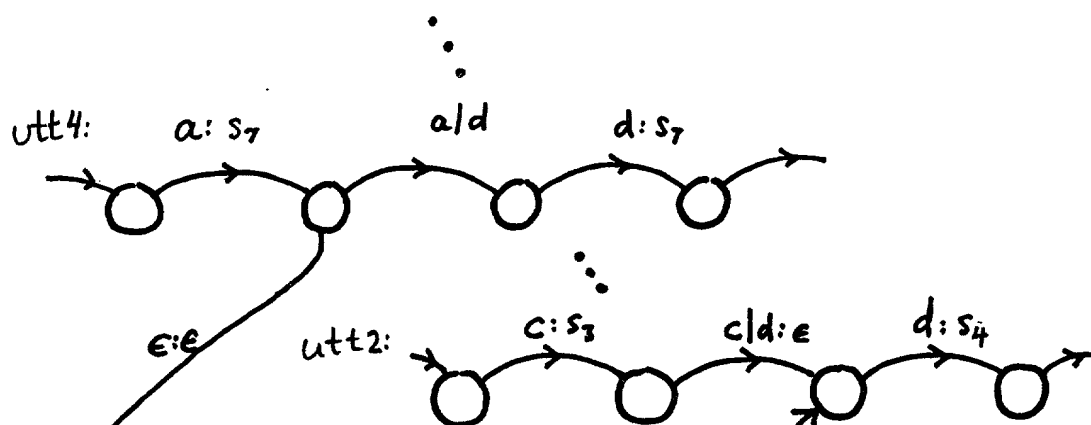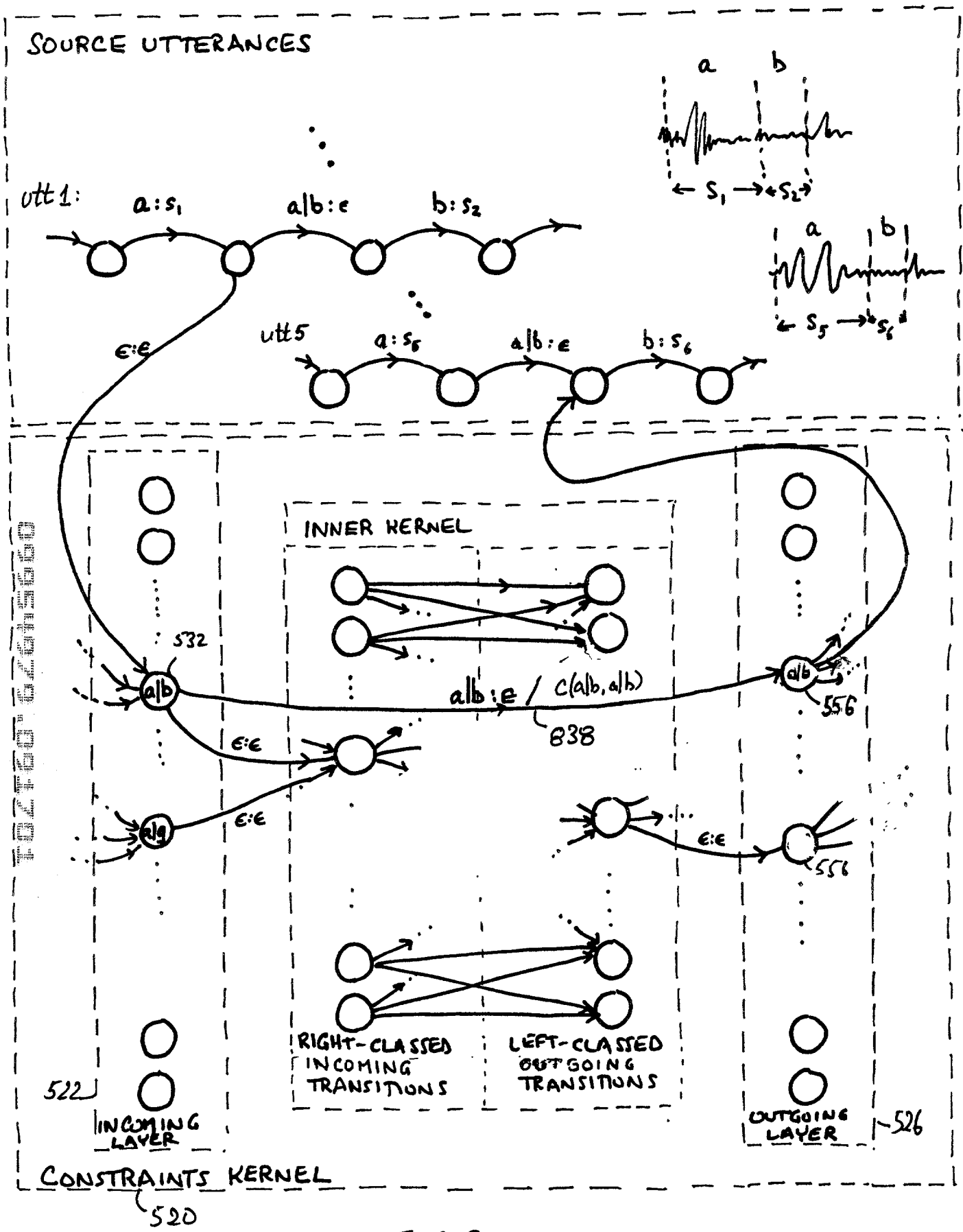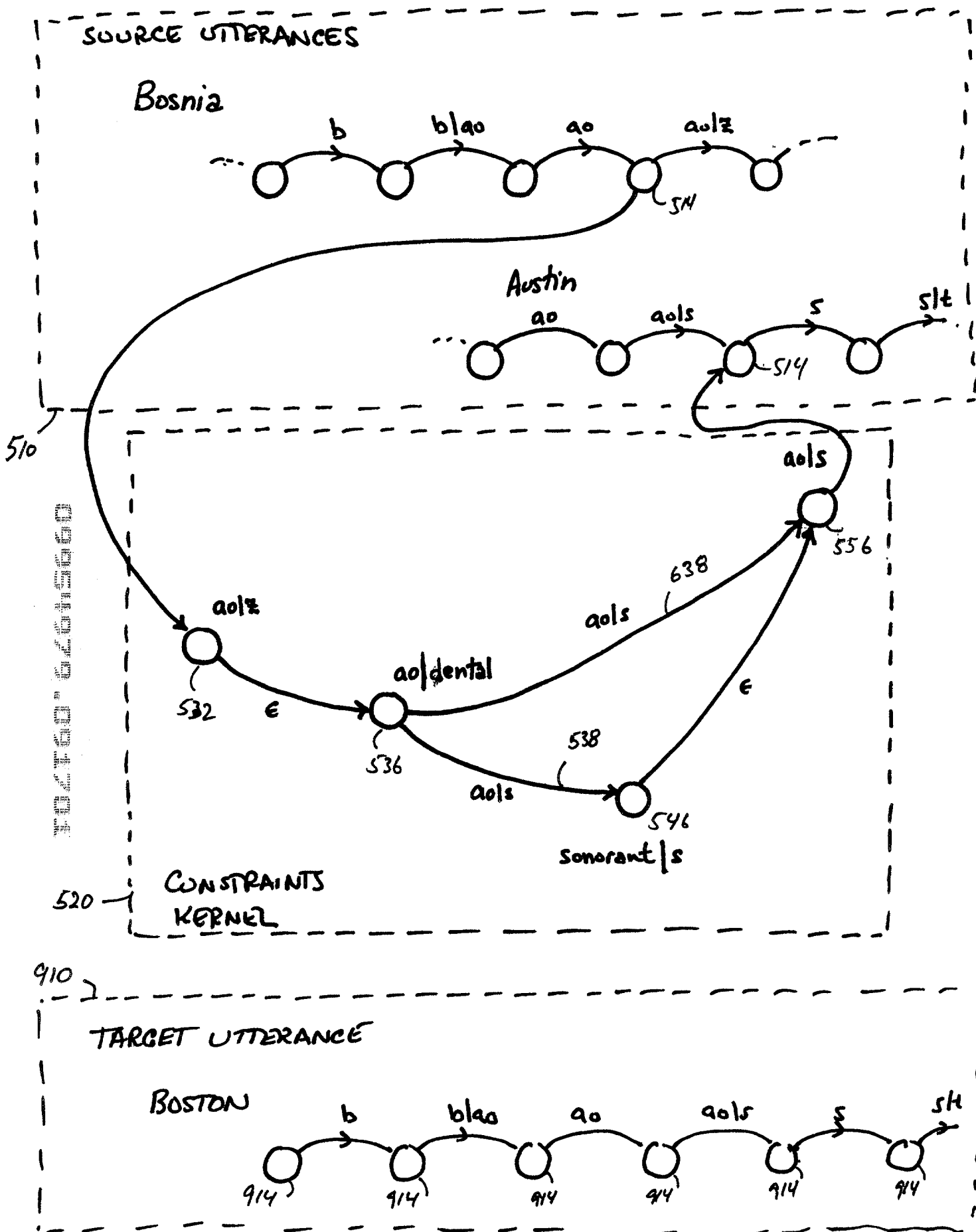| vowel:	y iy ih ix ux ey eh ae ay oy uw uh ah ax ow er aa ao aw
semivowel:	w l r el
nasal:	m em n en ng
aspirant:	hh
obstruent:	b bcl p pcl f v d dcl t tcl s z sh zh ch jh th dh dx g gcl k kcl
silence:	h# pau

FIG. 10A

	vowel	semivowel	nasal	obstruent	aspirant	silence
vowel	10000	10000	7500	10	10	10
semivowel	1000	7500	7500	10	10	10
nasal	5000	10	10	10	10	10
obstruent	10	10	10	10000	10	10
aspirant	5000	5000	10	5000	10	10
silence	10	10	10	10	10	10

FIG. 10B

vowel: y iy ih ix ux ey eh ae ay oy uw uh ah ax ow er aa ao aw w l r el
fricative: f v s z th dh dx sh zh ch jh hh
stop: b bcl d dcl g p pcl t tcl k kcl
nasal: m em n en ng
silence: h# pau

FIG. 11

labial: b bcl p pcl f v w
dental: d dcl t tcl s z sh zh ch jh th dh dx
velar: g gcl k kcl
nasal_labial: m em
nasal_dental: n en
nasal_velar: ng
front: y iy ih ix ux ey eh ae ay oy
back: uw uh ah ax ow er aa ao aw l r el
none: hh h# pau

FIG. 12A

	labial	alveolar	velar	m	n	ng	front	back	none
labial	10	1000	1000	1000	1000	1000	1000	1000	1000
alveolar	1000	10	1000	1000	1000	1000	1000	1000	1000
velar	1000	1000	10	1000	1000	1000	1000	1000	1000
m	100	1000	1000	10	1000	1000	1000	1000	1000
n	1000	100	1000	1000	10	1000	1000	1000	1000
ng	1000	1000	100	1000	1000	10	1000	1000	1000
front	1000	1000	1000	1000	1000	1000	10	1000	1000
back	1000	1000	1000	1000	1000	1000	1000	10	1000
none	1000	1000	1000	1000	1000	1000	1000	1000	10

FIG. 12B

	labial	alveolar	velar	m	n	ng	front	back	none
labial	10	1000	1000	1000	1000	1000	1000	1000	1000
alveolar	1000	10	1000	1000	1000	1000	1000	1000	1000
velar	1000	1000	10	1000	1000	1000	1000	1000	1000
m	100	1000	1000	10	1000	1000	1000	1000	1000
n	1000	100	1000	1000	10	1000	1000	1000	1000
ng	1000	1000	100	1000	1000	10	1000	1000	1000
front	1000	1000	1000	1000	1000	1000	10	1000	1000
back	1000	1000	1000	1000	1000	1000	1000	10	1000
none	1000	1000	1000	1000	1000	1000	1000	1000	10

FIG. 12C

retroflex: r er
round: uw ux w ow aw
sonorant: y iy ih ix ey eh ae ay oy uh ah ax aa ao l el m em n en ng
other: b bcl p pcl f v d dcl t tcl s z sh zh ch jh th dh dx g gcl k kcl hh h# pau

FIG. 13A

retroflex: r er
round: uw ux w b bcl p pcl f v m em
sonorant: y iy ih ix ey eh ae ay oy aw uh ah ax ow aa ao l el n en ng
other: d dcl t tcl s z sh zh ch jh th dh dx g gcl k kcl hh h# pau

FIG. 13B

retroflex	10	100	100	100
round	100	10	100	100
sonorant	100	100	10	100
other	100	100	100	10

FIG. 13C

retroflex	10	100	100	100
round	100	10	100	100
sonorant	100	100	10	100
other	100	100	100	10

FIG. 13D

front:	y iy ih ix ey eh ae ay oy
back:	uh ah ax aa ao l el
retroflex:	r er
round:	uw ux w ow aw
other:	b bcl p pcl f v d dcl t tcl s z sh zh ch jh th dh dx g gcl k kcl m em n en ng hh h# pau

FIG. 14A

front:	y iy ih ix ey eh ae aw
back:	uh ah ow aa ao ay oy l el
retroflex:	r er
round:	uw ux w
schwa:	ax
other:	b bcl p pcl f v d dcl t tcl s z sh zh ch jh th dh dx g gcl k kcl m em n en ng hh h# pau

FIG. 14B

	front	back	retroflex	round	other
front	10	10	10	10	10
back	10	10	10	10	10
retroflex	10	10	10	10	10
round	10	10	10	10	10
other	500	500	500	500	10

FIG. 14C

	front	back	retroflex	round	schwa	other
front	10	100	100	100	500	100
back	100	10	100	100	500	100
retroflex	100	100	10	100	500	100
round	100	100	100	10	500	100
schwa	500	500	500	500	10	500
other	100	100	100	100	500	10

FIG. 14D

obstruent: b bcl p pcl f v d dcl t tcl s z sh zh ch jh th dh dx g gcl k kcl hh h# pau
sonorant: y iy ih ix ux ey eh ae aw uw uh ah ax ow er aa ao ay oy w l r el m em n en ng

FIG. 15A

voiced: b bcl v d dcl z zh jh dh g gcl dx h# pau
unvoiced: p pcl f t tcl s sh ch th k kcl hh
sonorant: y iy ih ix ux ey eh ae aw uw uh ah ax ow er aa ao ay oy w l r el m em n en ng
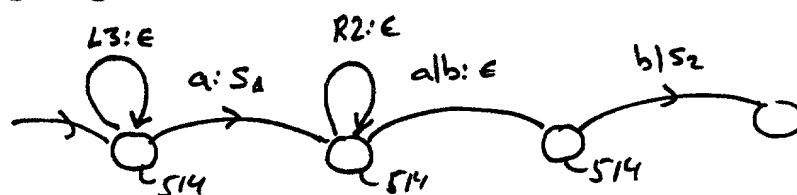
FIG. 15B

	obstruent	sonorant
obstruent	10	1000
sonorant	1000	10

FIG. 15C

	voiced	unvoiced	sonorant
voiced	10	100	1000
unvoiced	100	10	1000
sonorant	1000	1000	10

FIG. 15D

Source utterance 512

L3:ε   R2:ε
        a:s₁   a|b:ε   b|s₂
   ◯   ◯   ◯   ◯
    514   514   514

target utterance 912

        L1:L1/0   R3:R1/2   a|b:a|b
        L1:L2/1   a:a   R3:R2/1   ...
··· → ◯   ◯   →   ◯   ◯   →   ◯
        L1:L3/2   R3:R3/0
    914   914   914   914   914

FIG. 16